JONES DAY 3.0-001 FWC

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re patent application of: Gautvik, et al	Art Unit
Serial No.	
Filed: Herewith	Examiner
For: PRODUCTION OF HUMAN PARATHYROID HORMONE FROM MICROORGANISMS	

Hon. Comm. Patents & Trademarks
U.S. Patent & Trademark Office
Washington, D.C. 20231

Considered
1/21/90
RTeskin

DECLARATION OF DR. JANET KURJAN

1. I, JANET KURJAN, reside at Morningside Heights, New York City, New York. I am not an inventor of the above-identified patent application nor do I have any financial interest in such application. I have been asked by the inventors herein to assist technically in the prosecution of this application.

2. I have enclosed my curriculum vitae (Exhibit 1 hereto), which among other things, indicates my expertise and experience in the field of yeast genetics and, especially, in the field of use of the yeast mating factor including its use in genetic engineering. In this regard, I hold a Ph.D. degree from the University of Washington and I am currently a professor in the Department of Biological Sciences, Columbia University, New York, New York. I am also the same Janet Kurjan who is a coinventor of U.S. Patent No. 4,546,082, cited in this prosecution history, and who is an author of the article entitled "Structure of a Yeast Pheromone Gene (MFα):A Putative α-Factor Precursor Contains for Tandem Copies of

Mature ∝-Factor" published in Cell Volume 30, 933-943 (1982) (Exhibit 2 hereto).

3. My work in the yeast genetic field prior to the effective filing date of the above-identified application has extensively included the study of yeast pheromone gene for mating factor alpha as exemplified by my U.S. patent and the aforementioned publication. In particular, my patent teaches that the mating factor alpha pheromone gene could provide a convenient tool to molecular biologists in the expression, processing and secretion of heterologous genes as exemplified in my patent by somatostatin, enkaphelin and ACTH. This is the extent of the teachings of my work, that is, that mating factor alpha could be used as a leader sequence to assist in the facilitation of the expression and secretion of particular heterologous peptides.

4. However, this is not to say that this tool would always work or that all known mammalian peptides can effectively be expressed, processed and secreted using the mating factor alpha system. To my general knowledge, others have reported a lack of success using the mating factor alpha gene in the expression, processing and secretion of heterologous peptides. In this regard, I note that yeast signal sequences do not automatically substitute for mammalian signal sequences in expression of mammalian genes in yeast cells. Indeed, I was quite surprised to read that the Patent Office had attributed to my '082 patent a general teaching "applicable to obtaining secretion of heterologous polypeptides in Saccharomcyces [sic]". It is my opinion that this overstates the teaching of my patent.

5. My patent teaches that mating factor alpha can be a tool for expression, processing and secretion of heterologous peptides but not that it is a method applicable to all heterologous peptide expression in yeast (Saccaromyces cerevisiae). Indeed, as aforediscussed, I have reason to believe that the teaching of my invention is not broadly applicable to all heterologous polypeptides. One reason for reaching this opinion is that the expression of heterologous peptides in yeast systems is not predictable. Each system (and each hormone) must be evaluated on its own and tested on its own. This is precisely what the inventors of the above-identified application have accomplished in establishing the successful creation of a yeast expression system capable of expressing, processing and secreting human parathyroid hormone in an intact or undegraded state in which the hormone is biologically active.

6. I have carefully reviewed the above-identified patent application and all the prior art cited in the Patent Office Action of February 14, 1989. In so doing, I believe that the scientific work disclosed in the application to which this declaration refers is a noteworthy advance in the field of yeast genetics and not at all apparent from my teachings nor the teachings of Mahoney and Kronenberg et al. The teaching of Mahoney uses a very different genetic construction than was used in the current invention. Nothing about the Mahoney work suggests the construction used by the present inventors to achieve intact expression, processing and secretion of mature human parathyroid hormone. Kronenberg et al specifically state that they failed to achieve proper processing or secretion of

mature parathyroid hormone. Neither Mahoney nor Kronenberg et al uses the mating factor alpha gene. Therefore, the genetic constructions used by Mahoney and Kronenberg et al are not analogous to the genetic construction used in the current invention.

7. This work is made even more noteworthy by the known fragility of the human parathyroid hormone especially to protease. It is quite possible that human parathyroid hormone degradation occurs at a dibasic amino acid sequence rendering it susceptible to cleavage by the KEX-2 endopeptidase, the very protease that is responsible for successfully cleaving the pro region of mating factor alpha. The ability to exploit the mating factor alpha system, including exploiting the naturally occurring yeast proteolytic activity exemplified by the aforementioned KEX-2 endopeptidase is an accomplishment that I find admirable and for which the inventors of the above-identified application should be given recognition.

8. Similarly, it is a noteworthy aspect of this invention that despite the aforementioned difficulties, the inventors were able to overcome the noted problems of mating factor alpha when linked to a heterologous polypeptide, that is, with respect to the processing and secretion of the peptide. This the inventors have done, without any indication or teaching in the art cited in the Patent Office Action of February 14, 1989, including, in particular, my '082 patent.

9. I declare that all statements herein of my own knowledge are true and that all statements made on information and belief are believed to be true. I further state that these statements were made with knowledge that willful false

statements and the like so made are punishable by fine or imprisonment pursuant to appropriate sections of the United States Code and that such willful statements may jeopardize the validity of the above-identified patent application or any patent issuing therefrom.

Dated: August 3, 1989

Janet Kurjan, Ph.D.

CURRICULUM VITAE

NAME: Janet Kurjan

TITLE: Assistant Professor
Department of Biological Sciences
Columbia University (1982-present)

ADDRESS: Department of Biological Sciences
710A Sherman Fairchild Building
Columbia University
New York, NY 10027
(212) 854-8294

BIRTHDATE and PLACE: December 22, 1952; Flushing, New York

NATIONALITY: U.S.A.

EDUCATION:

1974--A.B., Biology
University of Chicago
Chicago, IL

1979--Ph.D., Genetics
Thesis advisor-Benjamin D. Hall
University of Washington
Seattle, WA

1980-1982--postdoctoral fellow
Postdoctoral advisor-Ira Herskowitz
University of Oregon
Eugene, OR and
University of California, San Francisco
San Francisco, CA

HONORS: 1980-1981--Damon Runyon-Walter Winchell Cancer Fund-Postdoctoral Fellowship

PROFESSIONAL ACTIVITIES:

Adhoc member Biomedical Sciences Study Section for postdoctoral fellowships, September 1987, July 1988.

Advisory committee member, American Chemical Society, Biotechnology Buyer's Guide, 1988-1989.

Co-editor, Signal Transduction in Invertebrates, Academic Press, (in planning stages).

FUNDING:

NIH 1 R01 GM40585
Pheromone response and G proteins in yeast
4/1/88 - 3/31/93
Direct costs (5 years)-$816,159

TEACHING EXPERIENCE:

1982-1985
C1005-Introduction to molecular and cellular biology, I
One-half semester, basic genetics, introductory level course

1983-present
G4064-Eukaryotic molecular genetics
Graduate level-selected topics in current molecular biology
2/3 lecture - 1/3 student presentations

1988-present
W3036-Applications of biotechnology
Small upper level undergraduate course
Application of recombinant DNA technology in medicine, agriculture, forensics, and the environment

MEETING & COURSE INVITATIONS (1988-1990):

Mating Type Control in Lower Eucaryotes
University of Wisconsin, Madison,
May, 1988
Cold Spring Harbor Symposium
Signal Transduction
May, 1988
Cold Spring Harbor Yeast Course
July, 1988
EMBO Guanine Nucleotide Binding Protein Workshop
The Netherlands
August, 1988
UNIDO-Molecular Genetics of Yeast Workshop
Trieste, Italy
March, 1989
Gordon Conference
Second Messengers and Protein Phosphorylation
June, 1989
Genetics Society
Co-plenary session GSA and Yeast Meetings
June, 1989
Cold Spring Harbor Yeast Course
August, 1989
Gordon Conference
Sensory Transduction in Microorganisms
January, 1990
ICN-UCLA Meeting
Guanine nucleotide-binding proteins
Feb., 1990

DOCTORAL DISSERTATION:

Fine structure analysis of the SUP4 locus in Saccharomyces cerevisiae, sponsor-Benjamin D. Hall, Department of Genetics, University of Washington, 1979.

PUBLICATIONS:

Olson, M.V., Page, G.S., Santenac, A., Loughney, K., Kurjan, J., Benditt, J. and Hall, B.D. (1980) Yeast suppressor tRNA genes, in Transfer RNA: Biological Aspects, Soll, D., Abelson, J. and Schimmel, P.R., eds. (New York, Cold Spring Harbor Laboratory), pp. 267-279.

Kurjan, J., Hall, B.D., Gillam, S. and Smith, M. (1980) Mutations at the SUP4 $tRNA^{Tyr}$ locus: DNA sequence changes in mutants lacking suppressor activity, Cell 20: 701-709.

Koski, R.A., Clarkson, S.G., Kurjan, J., Hall, B.D. and Smith, M. (1980) Mutations at the yeast SUP4 $tRNA^{Tyr}$ locus: Transcription of the mutant genes in vitro, Cell 22: 415-425.

Kurjan, J., Hall, B.D., Koski, R.A., Clarkson, S.G., Gillam, S. and Smith, M. (1981) Mutations at the yeast SUP4 $tRNA^{Tyr}$ locus: DNA sequence changes which affect transcription, in Molecular Genetics in Yeast; Alfred Benzon Symposium 16, Von Wettstein, D., Friis, J., Kielland-brandt, M. and Stenderup, A., eds. (Munksgaard, Copenhagen), pp 245-263.

Hopper, A.K., Nolan, S.L., Kurjan, J. and Hama-Furukawa, A. (1981) Genetic and biochemical approaches to studying in vivo intermediates in tRNA biosynthesis, in Molecular Genetics in Yeast; Alfred Benzon Symposium 16, Von Wettstein, D., Friis, J., Kielland-brandt, M. and Stenderup, A., eds. (Munksgaard, Copenhagen), pp. 302-325.

Koski, R., Clarkson, S., Kurjan, J., Hall, B., Gillam, S. and Smith, M. (1981) Transcription initiation and termination signals in the yeast SUP4 $tRNA^{Tyr}$ gene, in Developmental Biology Using Purified Genes: ICN-UCLA Symposium on Developmental Biology Using Purified Genes, Brown, D.D., ed. (Academic Press, New York), pp. 473-482.

Hopper, A.K. and Kurjan, J. (1981) tRNA synthesis: Analysis of in vivo precursor tRNAs from parental and mutant yeast strains, Nucl. Acids Res. 9: 1019-1029.

Kurjan, J. and Hall, B.D. (1982) Mutations at the yeast SUP4 $tRNA^{Tyr}$ locus: Isolation, genetic fine-structure mapping, and correlation with physical structure, Mol. Cell. Biol. 2: 1501-1513.

Nishikura, K., Kurjan, J., Hall, B.D. and DeRobertis, E.M. (1982) Genetic analysis of the processing of a spliced tRNA, EMBO J. 1: 263-268.

Kurjan, J. and Herskowitz, I. (1982) Structure of a yeast pheromone gene (MFα): A putative α-factor precursor contains four tandem copies of mature α-factor, Cell 30: 933-943.

Kurjan, J. (1985) α-Factor structural gene mutations in Saccharomyces cerevisiae: Effects on α-factor production and mating, Mol. Cell. Biol. 5: 787-796.

Kurjan, J. and Lipke, P.N. (1986) Agglutination and mating activity of the MFα2-encoded α-factor analog in Saccharomyces cerevisiae, J. Bacteriol. 168: 1472-1475.

Kurjan, J. (1987) Yeast α-factor genes, in Genes Encoding Hormones and Regulatory Peptides, Habener, J.F. ed. (Humana Press, New Jersey), pp. 413-434.

Dietzel, C. and Kurjan, J. (1987) Pheromonal regulation and sequence of the Saccharomyces cerevisiae SST2 gene: A model for desensitization to pheromone, Mol. Cell. Biol. 7: 4169-4177.

Dietzel, C. and Kurjan, J. (1987) The yeast SCG1 gene: A G_α-like protein implicated in the a- and α-factor response pathway, Cell 50: 1001-1010.

Kurjan, J. and Dietzel, C., (1988) Analysis of the role of SCG1, a G_α homolog, and SST2 in pheromone response and desensitization in yeast, Cold Spring Harbor Symp. Quant. Biol. 53: 577-583.

Lipke, P.N., Wojciechowicz, D., and Kurjan, J., (1989) AGα1 is the structural gene for the Saccharomyces cerevisiae α-agglutinin, a cell surface glycoprotein involved in cell-cell interactions during mating, Mol. Cell. Biol., in press (August).

Kurjan, J., G proteins in the yeast, Saccharomyces cerevisiae, in G Proteins, eds. L. Birnbaumer and R. Iyengar, (Academic Press, New York), in press.

Kurjan, J. Role of a G protein homolog in yeast pheromone response, in EMBO Workshop Proceedings: The Guanine-Nucleotide Binding Proteins, NATO ASI Series (Plenum Publishing Corporation), in press.

In preparation

Caplan, S., Green, R., Rocco, J., and Kurjan, J. The role of glycosylation and structure of the MFα1 α-factor precursor in α-factor production and mating in yeast, in preparation.

Dietzel, C., Hirsch, J., and Kurjan, J. Site-directed mutations in SCG1, the yeast G protein involved in pheromone response: Role of the guanine nucleotide binding regions and the carboxy terminus in SCG1 function, in preparation.

Caplan, S., Green, R., and Kurjan, J. Comparative roles of MFα1 and MFα2 precursors to provide mating function, in preparation.

Kang, Y.-S., Kane, J., Tipper, D.J., Stadel, J., & Kurjan, J. Function of mammalian and hybrid mammalian/yeast G_α subunits in yeast pheromone response and mating, in preparation.

Book

Signal Transduction in Invertebrates (tentative title), R. Dottin & J. Kurjan, eds. (Academic Press), in preparation.

PATENT:

Kurjan, J. & Herskowitz, I. E. coli/Saccharomyces cerevisiae plasmid cloning vector containing the alpha-factor gene for secretion and processing of hybrid proteins, Patent Number: 4,546,082, Oct. 8, 1985.

PREVIOUS RESEARCH:

Undergraduate Research--1973-1974
Research Advisors-Shelley and Mike Esposito

My research involved a genetic analysis of a dominant temperature-sensitive sporulation mutation in yeast. There was a suggestion that the few spores produced by this mutant were aneuploid leading to the hypothesis that this sporulation mutation also resulted in a segregation defect. My results showed that the aneuploidy of the spores was due to triploidy of the original mutant.

Graduate Thesis Research-1975-1979
Thesis Advisor-Benjamin D. Hall

My thesis research involved the isolation and characterization of second-site mutations in the yeast SUP4 tRNATyr gene that eliminate suppressor function. The mutations were shown to be distributed throughout the region encoding the mature tRNA, with one mutation in the intervening sequence. Further studies have indicated that these mutations affect many different levels of expression, including splicing, transcription termination, and transcription initiation.

Postdoctoral Research-1980-1982
Postdoctoral Advisor-Ira Herskowitz

My postdoctoral research involved the cloning and characterization of the major structural gene (MFα1) for the pheromone α-factor, a thirteen amino acid peptide that is secreted by α cells of yeast. This gene was shown to encode a large precursor, including a signal sequence, a pro region containing three N-glycosylation sites, and four repeats of the α-factor sequence. Comparisons to characterized proteolytic processing sites in precursors of higher eukaryotes suggested a pathway for the processing of mature α-factor from this precursor. Results of others have shown this pathway to be correct.

CURRENT RESEARCH:

Haploid yeast cells of opposite mating types, a and α, can mate with one another to form the a/α diploid. Two types of cell-cell interactions occur during the mating process. One interaction, adhesion of cells of opposite mating type, is mediated by cell surface glycoproteins, the a- and α-agglutinins. The second interaction involves the secretion of and response to the peptide pheromones, a-and α-factor. The work in my lab is aimed at elucidating several aspects of these cell-cell interactions.

During my first few years at Columbia, I continued studies of the α-factor structural genes, MFα1 and MFα2, that had been initiated during my postdoctoral research. Disruption mutagenesis showed that MFα1 and MFα2 are the only α-factor structural genes, that at least one of the two MFα genes is required for α mating, and that MFα1 is responsible for the majority of α-factor production. An unusual phenotype associated with one particular disruption mutant, a partially dominant inhibition of mating, and the inability of exogenous α-factor to efficiently alleviate the mating defect of α mfα1 mfα2 strains, led to the hypothesis that the MFα1 precursor plays a role in mating in addition to its role in α-factor production. The MFα1 precursor consists of a signal sequence, a pro region containing three N-glycosylation sites, and four repeats of the α-factor sequence. In particular, we had proposed that the pro region might play a role in mating. We have done site-directed mutagenesis studies to test this hypothesis. Deletion of a large portion of the pro region does not significantly affect mating, suggesting that our hypothesis was incorrect. We believe that the lack of

alleviation of the mating defect by exogenous α-factor may indicate that an a cell only mates efficiently with an α cell that is actively secreting α-factor. We have also mutated the N-glycosylation sites in the pro region to investigate the role of glycosylation of the MFα1 precursor in α-factor processing and secretion. A mutant with all three glycosylation sites eliminated secretes less α-factor than the wild-type and builds up a precursor which has undergone signal sequence cleavage, but no other processing steps. These results indicate that N-glycosylation is important, but not essential, for secretion of α-factor. The precursor buildup suggests that lack of glycosylation results either in less efficient processing by the lys-arg endopeptidase responsible for an early processing step or in less efficient transport from the ER to the Golgi, where this processing step occurs. At this time, I do not plan continue study of the MFα genes and plan instead to concentrate on the projects described below.

Several years ago, Peter Lipke (Hunter College) and I initiated a collaboration to study a- and α-agglutinins using a genetic and molecular approach, which would complement the biochemical approach that is ongoing in Peter's lab. We have isolated a- and α-specific agglutination defective mutants, at least some of which should identify the a- and α-agglutinin structural genes, and are using these mutants to clone the genes by complementation. We have a single complementation group, agα1, that results in an α-specific agglutination defect, consistent with the biochemical results that indicate that the α-agglutinin consists of a single polypeptide. We have isolated a clone that complements agα1 and have sequenced a 6 kb fragment that is sufficient for complementation. This fragment contains an open reading frame of 650 amino acids that shows features predicted for the α-agglutinin. The size is consistant with Peter's estimate for the size of the α-agglutinin, the amino terminal region shows features similar to signal sequences of secreted proteins, and there are twelve possible N-glycosylation sites and a high proportion of serines and threonines as expected for a protein that is modified by both N- and O-linked carbohydrate. Peter's lab has shown that AGα1 does encode the α-agglutinin by showing that an α-agglutinin-specific antibody recognizes a fusion protein containing sequences from the AGα1 open reading frame isolated from E. coli. Further studies being initiated currently include an investigation the regions of the α-agglutinin involved in binding to a-agglutinin and the possibility that a phosphatidylinositol anchor is involved in localization of the α-agglutinin to the cell surface.

The a-agglutinin is composed of a core and a binding fragment and consists of 90% carbohydrate. Our a-agglutination defective mutants correspond to multiple complementation groups, consistent with multiple subunits. Most of the mutants fall into one complementation group, and we have recently isolated clones that complement these mutations. We have narrowed down the complementing region to a 3 kb fragment and are initiating sequencing of this fragment. The long-term goal of this project is to elucidate the features of the a- and α-agglutinins that are involved in the protein-protein interaction and therefore mediate the direct cell-cell interaction between mating a and α cells.

Another aspect of the cell-cell interactions involved in mating is the response to pheromone. Each of the pheromones (a- and α-factor) interacts with a mating type specific receptor present on the opposite mating type. The two pheromone-receptor interactions are interchangable, indicating that a common intracellular pathway is involved in both responses. Recently, we have been investigating the intracellular response and desensitization pathway. Our initial approach involved the isolation of a gene (SST2) that had been implicated in this pathway. Mutations in the SST2 gene result in cells that are hypersensitive to the opposite pheromone and are defective in recovery from pheromone arrest. We

cloned the SST2 gene by screening of a multicopy yeast clone bank in an a sst2 strain for transformants that showed resistance to α-factor. One plasmid obtained in this screen was shown to contain the SST2 gene. Sequence analysis suggests that SST2 encodes a hydrophilic protein of 698 amino acids with no significant homology to any protein in the current data bases. It is expressed constitutively at low basal levels in a and α, but not a/α, strains, and its level of expression is induced in a and α cells by exposure to the opposite pheromone. We propose that SST2 is involved in desensitization of cells to the opposite pheromone, and would like to learn the mechanism of this desensitization process.

During the process of cloning SST2, we identified several other plasmids that are able to suppress sst2 mutations. One of these plasmids contains a gene, SCG1, that encodes a 472 amino acid protein that shows a very high level of homology to the α subunits of transducin and other G proteins, suggesting that SCG1 is also a guanine nucleotide binding protein. Disruption of SCG1 results in a haploid-specific phenotype; scg1 cells produce very small colonies consisting of large, unbudded cells. The rat α_s subunit, expressed under the control of an inducible yeast promoter, is able to complement the scg1 growth defect, indicating functional as well as sequence conservation. It is interesting to note that the structure of the a- and α-factor receptors is quite similar to the structure of receptors whose signal is mediated by G proteins, including rhodopsin, the β-adrenergic receptor, and the muscarinic receptor. We propose that SCG1 mediates transduction of the signal from the activated receptor to an effector, although our model for the role of SCG1 differs from the model for the role of G_s and transducin in the β-adrenergic and phototransduction systems, respectively. We have constructed site-directed mutations in SCG1 to further investigate SCG1 function. One mutation that is predicted to affect guanine nucleotide binding, based on analogies to EF-Tu and ras, results in a phenotype similar to the scg1 disruption phenotype, suggesting that guanine nucleotide binding is important for SCG1 function. Another mutation, a C-terminal truncation of SCG1, results in a sterile phenotype, consistent with our model for SCG1 function and the model of Bourne that predicts that the C-terminus of G_α subunits is involved in the interaction with the receptor. Other mutations also result in interesting phenotypes. We plan to continue to study the role of SCG1 in pheromone response, and to identify additional components of this system. One high priority is to identify the effector in this system, which is currently unknown.

Cell, Vol. 30, 933–943, October 1982, Copyright © 1982 by MIT

Structure of a Yeast Pheromone Gene (*MFα*): A Putative α-Factor Precursor Contains Four Tandem Copies of Mature α-Factor

Janet Kurjan* and Ira Herskowitz†
Department of Biology
and Institute of Molecular Biology
University of Oregon
Eugene, Oregon 97403
and Department of Biochemistry and Biophysics
University of California, San Francisco
San Francisco, California 94143

Summary

We have cloned and sequenced a gene (*MFα*) coding for α-factor, a tridecapeptide mating factor secreted by yeast α cells. A plasmid carrying the *MFα* gene was identified by screening for production of α-factor by *matα2* mutants, which fail to secrete α-factor because of simultaneous synthesis and degradation of the factor. The cloned segment codes for four mature α-factors within a putative precursor of 165 amino acids. The putative precursor begins with a hydrophobic segment that presumably acts as a signal sequence for secretion. The next segment, of approximately 60 amino acids, contains three potential glycosylation sites. The carboxy-terminal half of the precursor contains four tandem copies of mature α-factor, each preceded by spacer peptides of six or eight amino acids (variations of Lys-Arg-Glu-Ala-Asp-Ala-Glu-Ala), which are hypothesized to contain proteolytic processing signals.

Introduction

Mating in yeast appears to be facilitated by oligopeptide pheromones (mating factors) that cause arrest of cells of opposite mating type in the G1 phase of the cell division cycle (reviewed by Manney et al., 1981). α cells produce α-factor, which has two forms, 13 and 12 L amino acids in length, the latter lacking the amino-terminal Trp residue of the tridecapeptide. a cells produce a-factor, which is 11 L amino acids in length and has two forms that differ in the sixth residue (Leu or Val) (Betz et al., 1981). Synthesis of α-factor requires cellular RNA and protein synthesis (Scherer et al., 1974). On the basis of these observations and by analogy with mammalian peptide hormones, it has been proposed that the yeast mating factors are derived by proteolytic processing from a larger precursor.

We wanted to clone the α-factor structural gene (*MFα*, "α mating factor") for several reasons: to determine the mechanism of synthesis of α-factor and, in particular, to determine why only α cells and not a or a/α cells secrete α-factor; and to assess the physiological role of α-factor in mating by isolating and analyzing mutants defective solely in α-factor synthesis. The most straightforward avenue for cloning *MFα*, by complementation of a yeast mutant defective in this gene, was not possible because mutants defective in the α-factor structural gene have not been identified. There are, however, several mutants that do not produce α-factor (MacKay and Manney, 1974; Leibowitz and Wickner, 1976; Sprague et al., 1981). An understanding of the defect in α-factor synthesis in one of these mutants (defective in the *MATα2* gene) provided a rationale for cloning the α-factor gene.

According to the α1-α2 hypothesis for control of cell type (Strathern et al., 1981), the α mating type locus (*MATα*) codes for two regulatory proteins: α1, which stimulates expression of unlinked α-specific genes (perhaps including the α-factor structural gene), and α2, which inhibits expression of unlinked a-specific genes (Figure 1). *matα2*⁻ mutants express certain phenotypes characteristic of a cells, notably mating with α cells and degradation of α-factor, which are ordinarily inhibited by α2 product (Strathern et al., 1981; Hicks and Herskowitz, 1976a). Inability of *matα2* mutants to secrete α-factor results from expression of the a-specific gene (*BAR1*) responsible for degradation of α-factor—*matα2* mutants defective in *BAR1* produce α-factor (Sprague and Herskowitz, 1981). *matα2* mutants thus produce α-factor but degrade it at a rate sufficient to prevent its secretion. Our scheme for cloning the α-factor structural gene was based on the hope that *matα2* cells containing an α-factor structural gene on a high copy number plasmid would produce more α-factor than could be degraded. Such a plasmid was obtained and was shown to encode α-factor.

Results

Isolation of a Plasmid That Allows α-Factor Production by *matα2* Mutants

We have used a clone bank (constructed by K. Nasmyth) that contains random genomic fragments of yeast DNA inserted into high copy number plasmid YEp13 (Broach et al., 1979). This plasmid contains the origin of replication of the yeast 2μ plasmid and is present in 30–50 copies per yeast cell (Broach and Hicks, 1980; J. Broach, personal communication). A *matα2 leu2* recipient (strain XK41-10b) was transformed with plasmid DNA isolated from the clone bank, and 17,000 Leu⁺ transformants were selected, pooled and replated on selective medium (lacking leucine). Approximately 50,000 colonies were screened for production of α-factor by the halo method (Figure 2; see Experimental Procedures). One colony (K69) exhibited a small, distinct halo on lawns of a tester cells. To confirm that the halo resulted from production of α-factor and not, for example, from production of killer toxin (which also forms a halo on

* Present address: Department of Biological Sciences, Columbia University, New York, New York 10027.
† Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143.

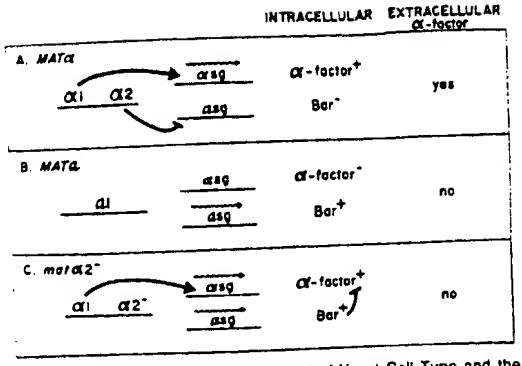

Figure 1. The α1-α2 Model for Control of Yeast Cell Type and the Behavior of *matα2* Mutants

The mating type locus is shown on the left for an α cell (A), an a cell (B) and a *matα2* mutant (C). *asg* and *esg* are α-specific and a-specific genes, respectively. Bold line with arrowhead indicates stimulation of gene activity; bold line with blunt end indicates inhibition of synthesis (for *asg*) or activity (for α-factor).

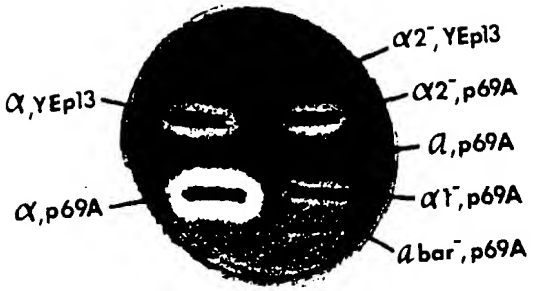

Figure 2. α-Factor Production by Yeast Strains Carrying Plasmids p69A or YEp13

α-Factor production was determined by the halo assay on a lawn of *MATa sst2* strain RC631 (see Experimental Procedures). Strains are (clockwise from the upper right): [YEp13]/*matα2-4*, [p69A]/*matα2-4*, [p69A]/*MATa*, [p69A]/*matα1-5*, [p69A]/*MATa bar1-1*, [p69A]/*MATα* and [YEp13]/*MATα*. Recipient strains carrying plasmids were XK41-10b (*matα2*), DC5 (*MATa*), XK96A2-6b (*matα1-5*), DC6 (*MATα*) and G245-24C (*MATa bar1*).

sensitive hosts; Fink and Styles, 1972), we assayed K69 for α-factor by the confrontation assay (see Experimental Procedures). K69 secreted a factor that elicited cell cycle arrest and aberrant cell morphology in a cells, indicating that it produces α-factor. Production of α-factor by K69 requires presence of a plasmid (p69A) in these cells: colonies grown nonselectively (in the presence of leucine) that lose the Leu+ plasmid also lose halo-forming ability. Plasmid p69A thus appears to contain a determinant that allows the *matα2* recipient to produce α-factor.

Behavior of Plasmid p69A in Different Strains

Plasmid p69A was introduced into various yeast strains differing in their mating type locus and in the *BAR1* gene to determine whether it affected α-factor synthesis (Table 1; Figure 2). α-Factor was assayed by halo formation on lawns of the most sensitive a tester strain (RC631), which carries the *sst2-1* mutation (Chan and Otte, 1982). As expected, reintroduction of p69A into a *matα2-4* strain yielded cells that produced α-factor. This strain is still defective in mating; hence, p69A does not simply contain a functional *MATα2* gene. An important expectation for a high copy number plasmid carrying an α-factor structural gene is that it should result in overproduction of α-factor in a wild-type *MATα* strain. Plasmid 69A shows this property, as indicated by an increase in halo size (compared to a [YEp13]/*MATα* strain) when tested on a *MATa sst1-2* lawn. Similarly, [p69A]/*MATα* strains produce a small halo on a *MATa SST* lawn, whereas [YEp13]/*MATα* strains do not. Plasmid p69A thus contains a gene whose expression is limiting for α-factor production in wild-type α cells as well as in *matα2* mutants.

The properties of p69A in several additional strains are also consistent 1with the presence of an α-factor structural gene on this plasmid. [p69A]/*matα1-5* cells produce a measurable amount of α-factor (a small halo) but considerably less than [p69A]/*matα2-4* cells. The poor production of α-factor by [p69A]/*matα1-5* cells is not surprising since α1 product is necessary for α-factor synthesis. Whether α1 product is necessary for synthesis of α-factor RNA or, for example, synthesis of an essential processing enzyme is not known. The low level of α-factor produced by [p69A]/*matα1-5* cells might result from transcription initiating from adjacent plasmid sequences or from leakiness of the *matα1-5* mutation.

Plasmid p69A allows *MATa* strains defective in *BAR1* to produce a low level of α-factor. Halo sizes in different transformants vary but are in some cases as large as those produced by [p69A]/*matα1-5* cells. The ability of [p69A]/*MATa bar1* cells to produce some α-factor may reflect low-level α1-independent expression of the plasmid gene conferring α-factor production or α1-independent expression of another gene necessary for α-factor synthesis. [p69A]/*MATa BAR1+* cells do not produce α-factor. We presume that this strain produces α-factor at the same level as the [p69A]/*MATa bar1* strain, but that the α-factor is degraded by the *BAR1* gene product.

The ability of p69A to allow production of α-factor by *matα2* mutants can be explained in a variety of ways; for example, this plasmid may carry an α-factor structural gene or an antagonist of the *BAR1* gene product. Nucleotide sequence analysis described below shows that p69A codes for α-factor.

Plasmid p69A Codes for α-Factor and a Putative α-Factor Precursor

As a prelude to sequencing, the determinant for α-factor synthesis was localized within the 4 kb insert in plasmid p69A. The insert contains three Eco RI sites, which define segments R1-1, R1-2, R1-3 and R1-4 (see Figure 3). The R1-2 fragment contains a cluster

Table 1. Production of α-Factor by Strains Carrying Plasmid p69A

Plasmid	Nuclear Genotype	α-Factor Production Assayed on Tester Lawns	
		MATa sst2-1	*MATa SST2*
p69A	*matα2-4*	+	−
p69A	*MATα*	+++	+
p69A	*matα1-5*	+/−	−
YEp13	*MATα*	++	−
p69A	*MATa bar1-1*	+/−[a]	−
p69A	*MATa BAR1*[b]	−	−

Strains carrying various plasmids were tested for α-factor production by the halo assay at 30°C on BBSD medium (to select for maintenance of the plasmid) as described in the Experimental Procedures. *MATa* testers were strains 227 (*SST2*) and RC231 (*sst2-1*). (++) wild-type α-factor halo; (+) halo smaller than wild-type; (+++) halo larger than wild-type; (+/−) small halo; (−) no halo.

[a] Halo size varies in different segregants from no halo to a halo as large as by [p69A]/*matα1-5* strains.

[b] This [p69A]/*MATa* strain did not show increased level of a-factor (J. Kurjan, unpublished observations). *matα2-4*, *matα1-5* and *MATa* strains carrying YEp13 did not produce α-factor. Strains are described further in the legend to Figure 2.

of four Hind III sites, which are separated from each other by 63 bp. Plasmids lacking different Eco RI fragments were constructed by partial digestion with Eco RI followed by religation and were tested for α-factor production in XK41-10b. These results indicate that the α-factor determinant is located in the 1.7 kb fragment R1-2. All plasmids containing R1-2 allow α-factor production, whereas all plasmids lacking this fragment do not (Table 2). Plasmids that carry either the region of R1-2 to the left to the Hind III site cluster (YEp13-H2) or the region to the right of the cluster (YEp13-H1) do not produce α-factor (Figure 3). These results indicate that the Hind III sites of R1-2 lie within the gene responsible for α-factor production.

The sequencing strategy is summarized in Figure 4. First, Hind III sites 1 and 4 were 3′-end-labeled, and the sequences to the right of site 4 (and to the left of site 1) were determined. The last T of Hind III site 4 (position 457) is the start of a sequence that codes for the α-factor tridecapeptide, which is followed immediately by a TAA termination signal (Figure 5). Sequencing leftward from a Sal I site 40 bp downstream from the α-factor coding sequence revealed a surprise—the presence of four tandem sequences coding for α-factor. Each of these sequences is preceded by a similar coding sequence of 18 or 24 nucleotide pairs. Further sequence determination reveals an ATG located 89 amino acids upstream from the first α-factor sequence. We have sequenced both DNA strands between the Hinf I site 1 and Sal I, which includes the entire coding region and some flanking regions. The remainder of the sequence has been determined for only a single DNA strand. The sequence is shown in Figure 5.

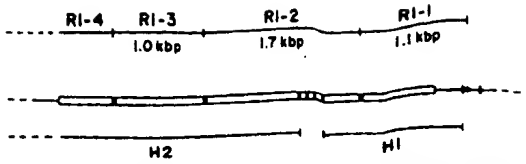

Figure 3. Partial Restriction Endonuclease Map of the Insert Present in p69A

Rectangle indicates the yeast DNA fragment inserted into the Bam HI site of YEp13; flanking regions are from YEp13. The inserted segment is drawn to approximate physical scale. Circles indicate Eco RI sites; vertical bars indicate Hind III sites. R1-1, R1-2, R1-3 and R1-4 are Eco RI fragments; H1 and H2 are Hind III fragments. α-Factor production by plasmid derivatives lacking various fragments is given in Table 2 and shows that the *MFα* gene is located in fragment R1-2.

Nucleotide sequence analysis thus reveals that plasmid p69A contains sequences coding for α-factor that are contained within a large coding sequence, a putative α-factor precursor. We summarize the key features of this sequence here and present further discussion below. First, the precursor is 165 amino acids long, beginning with ATG and ending at TAA immediately following α-factor sequence 4. Second, the precursor contains four sequences coding for identical α-factor tridecapeptides (α-factor 1, 2, 3, 4). Third, between these α-factor sequences are regions coding for similar octapeptides, which we refer to as peptide "spacers" S2, S3 and S4. Preceding α-factor 1 is a similar amino acid sequence of six amino acids (spacer 1, S1).

The Structure of the α-Factor Gene in Plasmid p69A Reflects the Genomic Sequence

Does the yeast genome contain a *MFα* gene with four tandem α-factor coding sequences? Because our screen for plasmids carrying an α-factor gene was based on overproduction of α-factor, it is conceivable that the α-factor coding sequences observed in plasmid p69A resulted from tandem amplification of one or more α-factor coding sequences during propagation of the plasmid in E. coli or in yeast. It is also possible that the genomic α-factor gene contains more than four copies of the α-factor coding sequence and that some copies are deleted during propagation of p69A in E. coli or in yeast. Several observations indicate that the structure of the insert in p69A reflects the α-factor gene in the yeast genome. First, all four α-factor coding sequences are not identical—α-factors 2 and 3 are identical, but differ from α-factor 1 at two positions and from α-factor 4 at five. It thus appears that the four α-factor coding sequences did not arise simply by quadruplication of a single genomic coding sequence. Second, genomic DNA sequences homologous to α-factor coding sequences of p69A have been isolated from a clone bank by hybridization (see Experimental Procedures) and thus without selection for overproduction of α-factor in yeast. The plasmids that have been analyzed contain an

Table 2. Localization of the α-Factor Determinant in p69A

YEp13 Derivative Containing Fragments	α-Factor Production by *matα2-4* Strain Carrying Plasmid
R1-1, R1-2, R1-3, R1-4	+
R1-1, R1-2, R1-4	+
R1-2, R1-3, R1-4	+
R1-4	−
R1-1, R1-4	−
R1-3, R1-4	−
R1-1, R1-3, R1-4	−
H1	−
H2	−

The plasmid carrying fragment H1 contains the H1 Hind III fragment from p69A inserted into YEp13. All other plasmids are deletion derivatives of p69A. Plasmid structures are shown in Figure 3. Plasmids were inserted into *matα2-4* strain XK41-10b and tested for α-factor production as described in the Experimental Procedures.

insert whose size is identical with that in p69A. Finally, and most definitively, the yeast genome contains a segment identical in size with the plasmid segment coding for α-factor, as shown in the following analysis. The four tandem copies of α-factor coded by p69A are on a Pst I–Sal I fragment of 510 bp. DNA from the yeast strain from which the clone bank was constructed (AB320) and from plasmid p69A were cleaved with Pst I and Sal I and fractionated on high percentage agarose gels, on which a difference in size of one repeat unit (63 bp, spacer plus α-factor) was clearly distinguishable. The DNA fragments were then probed with a mixture of end-labeled 63 bp Hind III fragments (corresponding to α-factor 1 through spacer 4). As shown in Figure 6, the probe hybridizes to identical fragments of approximately 500 bp that are present in both AB320 and p69A digests. The yeast genome thus appears to have the same size α-factor coding fragment as the plasmid and thus presumably contains four α-factor repeats.

A similar hybridization analysis of AB320 DNA cleaved with Hind III indicates that there are additional nucleic acid sequences that exhibit weak hybridization to α-factor probes and thus contain some homology to α-factor (J. Kurjan, unpublished observations). One segment contains homology only to the spacer. These sequences are under study.

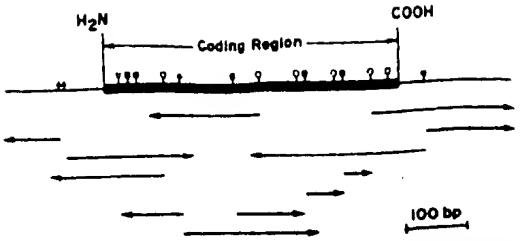

Figure 4. Strategy for DNA Sequencing and Positions of Restriction Endonuclease Sites

Arrows indicate direction and extent of sequence determination from the indicated sites. Restriction endonuclease sites: (▽) Pst I; (▲) Hinf I; (△) Taq I; (○) Hind III; (●) Sal I; (□) Hpa II; (■) Fnu 4H. Hinf I sites 1 and 2 correspond to the right and left Hinf I sites, respectively.

Discussion

Cloning of an α-Factor Structural Gene

Many yeast genes have been cloned by screening for ability of plasmids to supply the missing gene product to a mutant yeast strain (Broach et al., 1979; Nasmyth and Reed, 1980). Our method for cloning an α-factor gene represents a variation and extension of this method in that we did not rely on a mutation in the structural gene that we wanted to clone. Rather, we screened for plasmids that allowed overproduction of α-factor to a level sufficient to escape from the degradation system expressed in *matα2* mutants. A confirmation that plasmid p69A overproduces α-factor comes from our observation that wild-type α cells carrying p69A overproduce α-factor (Figure 2). Quantitative assays show that α-factor activity is increased at least sixfold and that α-factor cross-reacting material is increased 30-fold (D. Julius and J. Thorner, personal communication), comparable to the level expected from an increase in α-factor due simply to increased dosage of the *MFα* gene carried on YEp13. Failure of cells carrying p69A to produce as much α-factor as cross-reacting material is presumably due to limitation in α-factor processing.

In principle, overproduction of α-factor might also occur if the α-factor gene is under control by plasmid regulatory signals. Although we cannot rule out this explanation for p69A, the inserted DNA contains ample information flanking the coding sequence (approximately 2.5 kb before the coding sequence and 1.3 kb after it) and presumably contains the normal reg-

Figure 5. Nucleotide Sequence of *MFα*

The gap between positions 250 and 260 is drawn to allow alignment of the sequence coding for α-factor and to maximize the nucleotide sequence homologies among the spacer sequences. Asterisks indicate spacer sequences; α-factor coding regions are to the right (beginning at amino acid positions 90, 111, 132 and 153). The overlined sequence at −128 to −122 is a putative TATA sequence; the overlined sequences beginning at positions 526 and 596 show homology with a consensus sequence at the 3′ end of yeast genes (see text). Potential glycosylation sites (beginning at amino acid positions 23, 57 and 67) are indicated by amino acids in capital letters.

Other open reading frames include one beginning at −97 and continuing rightward (in the same direction as *MFα*), through two UGA codons, for 92 amino acid residues; the longest leftward open reading frames begin at 660 and 116 and continue for 62 and 52 amino acids, respectively. α-Factor cannot be encoded by either DNA strand. Sequencing reactions showed gaps corresponding to the positions coding for Pro-Gly in α-factors 1, 2 and 3, apparently as a result of methylation of Eco RII sites (CC*TGG) beginning at positions 289, 353 and 415.

```
                                                                      -170
                                                                      AGTG

   -160     -150      -140      -130      -120      -110     -100      -90
CAAGAAAACCAAAAAGCAACAACAGGTTTTGGATAAGTACATATATAAGAGGGCCTTTTGTTCCCATCAAAAATGTTACTGTT

   -80      -70       -60       -50       -40       -30      -20       -10
CTTACGATTCATTTACGATTCAAGAATAGTTCAAACAAGAAGATTACAAACTATCAATTTCATACACAATATAAACGACCAAA

    1           10           20  PstI 30             40               50          60
AGA ATG AGA TTT CCT TCA ATT TTT ACT GCA GTT TTA TTC GCA GCA TCC TCC GCA TTA GCT GCT
    met arg phe pro ser ile phe thr ala val leu phe ala ala ser ser ala leu ala ala
    1                                   10                                      20

            70           80          90           100         110          120
CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC
pro val ASN THR THR thr glu asp glu thr ala gln ile pro ala glu ala val ile gly tyr
                                    30                                      40

        130          140          150          160          170          180
TCA GAT TTA GAA GGG GAT TTC GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG
ser asp leu glu gly asp phe asp val ala val leu pro phe ser ASN SER THR asn asn gly
                                50                                      60

    190          200          210          220          230          240
TTA TTG TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA TCT TTG GAT
leu leu phe ile ASN THR THR ile ala ser ile ala ala lys glu glu gly val ser leu asp
                            70                                      80

250                     260    HindIII 270         280          290          300
AAA AGA GAG             GCT GAA GCT TGG CAT TGG TTG CAA CTA AAA CCT GGC CAA CCA ATG TAC
lys arg glu             ala glu ala trp his trp leu gln leu lys pro gly gln pro met tyr
*********************************  90                                          102

  310          320          HindIII       340          350          360
AAG AGA GAA GCC GAA GCT GAA GCT TGG CAT TGG CTG CAA CTA AAG CCT GGC CAA CCA ATG TAC
lys arg glu ala glu ala glu ala trp his trp leu gln leu lys pro gly gln pro met tyr
*********************************  111                                        123

370          380          HindIII       400          410          420          430
AAA AGA GAA GCC GAC GCT GAA GCT TGG CAT TGG CTG CAA CTA AAG CCT GGC CAA CCA ATG TAC
lys arg glu ala asp ala glu ala trp his trp leu gln leu lys pro gly gln pro met tyr
*********************************  132                                        144

        440          450  HindIII 460          470          480          490
AAA AGA GAA GCC GAC GCT GAA GCT TGG CAT TGG TTG CAG TTA AAA CCC GGC CAA CCA ATG TAC
lys arg glu ala asp ala glu ala trp his trp leu gln leu lys pro gly gln pro met tyr
*********************************  153                                        165

     500       510       520       530  SalI 540      550       560       570
TAA GCCCGACTGATAACAACAGTGTAGATGTAACAAAGTCGACTTTGTTCCCACTGTACTTTTAGCTCGTACAAAATACAAT
stop

   580       590       600       610       620       630       640       650       660
ATACTTTTCATTTCTCCGTAAACAACCTGTTTTCCCATGTAATATCCTTTTCTATTTTTCGTTTCGTTACCAACTTTACACAT

       670
ACTTTATATAGCTAT
```

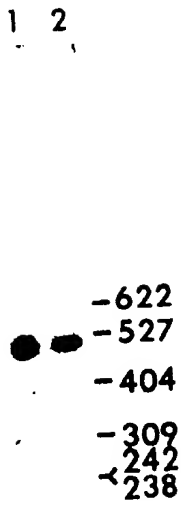

Figure 6. Comparison of Genomic and Plasmid Pst I–Sal I Fragments Containing *MFα*

(Lane 1) p69A DNA; (lane 2) DNA from yeast strain AB320. DNA was digested with Pst I and Sal I, fractionated on 2% agarose gels, transferred to nitrocellulose filters and probed with a mixture of 63 bp Hind III fragments (which contain only α-factor and spacer coding sequences) as described in the Experimental Procedures and in the text. The size markers, pBR322 DNA cleaved with Hpa II, were run in parallel on the same gel.

ulatory signals. Hence the α-factor gene need not be expressed from plasmid regulatory signals. We find that production of α-factor by strains carrying p69A still requires the α1 product, but we do not know whether this requirement reflects activation of the *MFα* gene itself by α1 or activation of another gene necessary for α-factor synthesis.

It has been suspected for a long time that the α-factor structural gene is distinct from the α mating type locus, although the absence of mutations in the α-factor structural gene prohibited a conclusive demonstration. Nucleotide sequence analysis of *MATα* indicates that it does not code for α-factor (Astell et al., 1981; M. Smith, personal communication). Our analysis of the *MFα* gene confirms this point.

The most striking feature of the α-factor gene from a strictly genetic standpoint is that it has four direct repeats of a closely homologous 63 nucleotide pair sequence: S2-αF2 (repeat 2) is homologous to S3-αF3 (repeat 3) in 61 of 63 positions and to S4-αF4 in 56 of 63 positions; repeats 3 and 4 are homologous in 58 of 63 positions. S1-αF1 differs from the other three repeats in lacking six nucleotide pairs in S1 but otherwise shows extensive homology. The longest direct repeat present in *MFα* is 63 nucleotide pairs, beginning at position 322 (or 323. . .331) and at position 385 (or 386. . .394). In other words, αF2-S3 and αF3-S4 are perfect 63 nucleotide pair direct repeats. Because our nucleotide sequence determination required propagation of the *MFα* gene in E. coli, we do not know whether heterology in the repeats has been reduced by homogenotization. In any event, it appears that the yeast genome contains four repeated sequences with extensive homology. Given this structure, one would expect that recombination among these repeats should lead to contraction and expansion of the α-factor gene to produce *MFα* genes with, for example, two or six repeats. The four repeats might be maintained by inhibiting recombination in this region or by selecting for maintenance of the repeats, for example, if a processing intermediate containing the four α-factor sequences performs some essential function.

The repeated structure of the *MFα* gene provides an explanation for some puzzling observations on α-factor-deficient mutants and leads to the proposal that the presence of a certain number of the repeats is required for α-factor production. Manney et al. (1981) isolated yeast mutants that produce reduced levels of α-factor and found that several of these mutants, when crossed to wild-type strains, yielded segregants with a continuum of levels of α-factor production. Although the wide variation of α-factor levels can be explained by segregation of several modifier loci present in the strains, a more interesting explanation is that these α-factor-deficient mutants have a reduced number of α-factor repeats and that this genetic defect can be repaired by amplifying the repeats during meiosis. Given our *MFα* probe, this hypothesis can be readily tested.

In addition to the 63 nucleotide pair repeats, the spacer regions contain a repeating unit of Glu-Ala or Asp-Ala. S1 contains two such pairs; S2, S3 and S4 contain three. It will be interesting to determine whether S. cerevisiae strains from diverse sources have the same number of these dipeptide repeats.

5′ and 3′ Regions and Codon Usage

The ATG codon (at position 1) is the only ATG codon that is in frame with the spacer and α-factor coding sequences and is thus the presumptive translation initiation codon for the α-factor precursor. This codon is preceded by in-phase stop codons at −15 and −60 and out-of-phase stop codons at −124 and −136. Two additional open reading frames are present in the 668 nucleotide region that precedes the α-factor precursor stop codon. One, at position 148, codes for a peptide of 17 amino acids. The other begins at −97 and codes for a polypeptide of 60 amino acids before a TGA codon. If we consider that TGA does not necessarily lead to translation termination (as suggested by the results of Astell et al., 1981), then a polypeptide of 92 amino acids can be produced (ig-

noring also a TGA codon corresponding to amino acid position 68). The significance of these other reading frames is unknown.

Most yeast genes that have been sequenced contain an A residue three nucleotides before the initiation codon; the *MFα* gene does too. Beginning 128 nucleotides before the presumptive initiation codon is a sequence (TATATAA) that closely resembles the TATA box found in several yeast genes (see Bennetzen and Hall, 1982a). If *MFα* proves to be positively regulated by the α1 product of the α mating type locus, it will be of interest to determine whether genes that are positively regulated by α1 (such as *STE3*; G. F. Sprague, Jr., R. Jensen and I. Herskowitz, manuscript submitted) share common sequences.

The region beyond the α-factor coding sequence lacks the sequence AATAAA that is found near the poly(A) addition site of most eucaryotic genes (Benoist et al., 1980). Yeast genes appear to have a conserved sequence beyond the coding sequence, which includes the consensus sequence TAAATAA$^{A}_{G}$ (Bennetzen and Hall, 1982a). The two sequences distal to the region coding for α-factor with the greatest homology to this sequence begin at position 526 (TAAcAAAG) and at position 596 (TAAAcAAc). This region does not contain a sequence similar to another consensus sequence (TAG. . . $^{TAGT}_{TATGT}$. . .TTT) that is found at the 3′ end of several yeast genes and that has been proposed to be involved in transcription termination (Zaret and Sherman, 1982).

Bennetzen and Hall (1982b) have noted that yeast has a strong preference in using 23 of the 59 nonunique sense codons. The *MFα* gene utilizes these preferred codons for 60% of its coding assigments (see Table 3); notable differences include codons for Leu, His and Gly.

Table 3. Codon Usage in the *MFα* Gene

Phe	UUU	4	Ser	UCU	1	Tyr	UAU	0	Cys	UGU	0
	UUC	2		UCC	3		UAC	5		UGC	0
Leu	UUA	5		UCA	2						
	UUG	5		UCG	0				Trp	UGG	8
Leu	CUU	0	Pro	CCU	4	His	CAU	4	Arg	CGU	0
	CUC	0		CCC	1		CAC	0		CGC	0
	CUA	3		CCA	6	Gln	CAA	8		CGA	0
	CUG	2		CCG	1		CAG	1		CGG	0
Ile	AUU	4	Thr	ACU	4	Asn	AAU	2	Ser	ACU	0
	AUC	1		ACC	0		AAC	3		AGC	2
	AUA	1		ACA	3	Lys	AAA	6	Arg	AGA	5
Met	AUG	5		ACG	1		AAG	3		AGG	0
Val	GUU	3	Ala	GCU	15	Asp	GAU	5	Gly	GGU	1
	GUC	2		GCC	4		GAC	2		GGC	4
	GUA	1		GCA	5	Glu	GAA	14		GGA	0
	GUG	0		GCG	0		GAG	1		GGG	3

Data are from the *MFα* sequence shown in Figure 5.

Expression of *MFα*

We do not know whether the α-factor gene that we have cloned is the gene used for α-factor production in wild-type yeast. A definitive answer to this question will come from replacement of the wild-type genomic sequence by mutant information contributed by the cloned gene. Several observations indicate, however, that the cloned α-factor gene is expressed in yeast. First, the cloned gene contains the appropriate punctuation signals for translation as well as putative signals for transcription and is expressed when carried on a plasmid. Second, *MATα* cells produce a poly(A) RNA species of approximately 500 nucleotides that hybridizes to the α-factor coding region of plasmid p69A, as determined by R-loop analysis carried out in conjunction with J. Yokum (data not shown). Finally, D. Julius and J. Thorner (personal communication) have found that an α strain defective in the *STE13* gene (Sprague et al., 1981) produces inactive forms of α-factor, some with four additional amino acids at the amino terminus of the mature factor. These four amino acids are those predicted from our sequence. In addition to the α-factor coding sequence that we have cloned, the yeast genome contains three sequences with homology to the α-factor coding sequence. Whether these sequences represent active α-factor genes is under study.

Putative α-Factor Precursor: Functional Domains and Processing

The *MFα* gene contains a coding sequence of 165 amino acids that we believe is an α-factor precursor (pre-pro-poly-α-factor). Because the coding sequence is uninterrupted and because R-loop analysis shows a continuous RNA–DNA duplex, there is no reason to invoke the existence of an intervening sequence in the *MFα* gene. The amino-terminal segment of 22 amino acids is hydrophobic and resembles signal sequences found in a wide variety of secretory protein precursors, which target polypeptides for processing and secretion pathways (Blobel and Dobberstein, 1975; Kreil, 1981). By analogy with other secretory proteins, we expect that the α-factor signal peptide is removed upon association with the endoplasmic reticulum, perhaps being cleaved after the alanine residue at position 19 or 20 (see Steiner et al., 1980), to yield pro-poly-α-factor (Figure 7). The signal peptide is followed by a (pro) segment of approximately 60 amino acids, which contains three potential glycosylation sites, Asn-X-Thr (Struck et al., 1978), at positions 23, 57 and 67. Although the function of this segment is unknown, one possibility is that it is involved in subsequent targetting of the α-factor precursor to other processing and secretion sites in the cell.

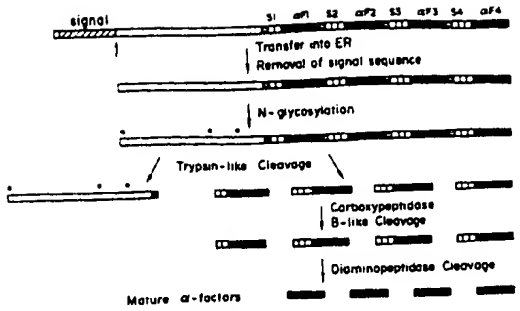

Figure 7. Scheme for Processing of the α-Factor Precursor

α-Factor oligopeptides are indicated by solid rectangles; the signal sequence by a diagonally crosshatched rectangle; spacer peptides consist of Lys-Arg (horizontally crosshatched box) followed by two or three Glu-Ala or Asp-Ala pairs (open boxes). The order of some of the steps (for example, involving cleavages by carboxypeptidase and dipeptidylaminopeptidase activities) is arbitrary. Other details are given in the text.

The second half of the α-factor precursor contains four tandem copies of mature α-factor, each preceded by a spacer peptide that presumably contains signals for processing. For many mammalian polypeptide hormones, pairs of basic amino acids flank the mature product, as in proinsulin (Steiner et al., 1980), proopiomelanocortin (Nakanishi et al., 1979) and proenkephalin (Stern et al., 1981; Comb et al., 1982). In these cases, action of a trypsin-like protease is thought to yield the mature hormone containing an additional amino acid (or two) at the carboxyl terminus, which can be removed by carboxypeptidase B action (see Hook et al., 1982). Processing of poly-α-factor may share common steps, but its processing must differ because the mature product is flanked by dibasic amino acids only at its carboxyl terminus. Cleavage of poly-α-factor with a trypsin-like protease would yield an intermediate with most of the spacer peptide linked to the amino terminus of α-factor and with Lys or Lys-Arg at its carboxyl terminus (see Figure 7). The carboxyl terminus could then be trimmed by a carboxypeptidase (such as yeast carboxypeptidase Y; Hayashi et al., 1973) or by a chymotryptic activity for α-factor 1, 2 and 3; no such processing is required for α-factor 4.

How is the correct amino terminus of α-factor produced? Although one could invoke an endopeptidase that removes the spacer segment by cleavage between Ala and Trp residues, findings on processing of honeybee promelittin lead to an attractive alternative hypothesis. Promelittin contains a 22 amino acid segment at the amino terminus of mature melittin, which includes the sequence *Glu-Ala-Glu-Ala-Asp-Ala-Glu-Ala*-Asp-Pro-*Glu-Ala* immediately preceding melittin. Kreil et al. (1980) have identified an enzyme, dipeptidylpeptidase IV, that sequentially removes the 11 amino acid pairs preceding melittin and stops at the amino terminus of melittin. G. Kreil (personal communication) has suggested that yeast cells may contain an analogous enzyme that is responsible for processing the amino terminus of the α-factor intermediate. Dipeptidylpeptidase activities are found in a wide variety of organisms, including yeast (Suarez Rendueles et al., 1981). Observations on the *ste13* mutant of yeast (D. Julius and J. Thorner, personal communication) may indicate that the *STE13* gene product is involved in such a processing reaction. Another yeast gene that seems likely to play a role in processing and secretion of α-factor is *KEX2*, which is required for production of the secreted killer toxin and also for α-factor (Leibowitz and Wickner, 1976).

Tanaka and Kita (1977) have described an oligopeptide secreted by α cells that contains 11 additional amino acids not attached to the amino terminus of α-factor and hence presumed to be linked at the carboxyl terminus. The amino acid composition of the additional segment shows no resemblance to the spacer peptide we have identified, nor can the additional segment be generated by translational readthrough of the stop codon at the end of the putative precursor. There are several possible explanations for the disagreement between our observations and those of Tanaka and Kita. The 24 amino acid form may be a bona fide precursor in some or all yeast strains, in which case the *MFα* gene we have cloned is used only in some strains or is not the only gene coding for α-factor. Another explanation is that the additional 11 amino acids may simply cochromatograph with mature α-factor rather than being covalently joined to it. It may be notable in this regard that nine of the 11 residues at positions 73–83 of the putative precursor match the additional amino acids found by Tanaka and Kita. Perhaps the precursor is processed to yield an oligopeptide containing residues 73–83 that has high affinity for mature α-factor.

The origin of the α-factor species containing only 12 amino acids is not illuminated by our studies. The dodecapeptide may be formed after maturation of the tridecapeptide or be formed from the proposed intermediate by improper cleavage, or it may be coded by a different *MFα* gene.

The finding that α-factor is produced as a precursor was no surprise. Since α-factor is a small, secreted oligopeptide, it seemed likely that an α-factor precursor would contain at least a signal sequence. The finding that the putative precursor contains four tandem copies of mature α-factor was, however, not anticipated. It appears that precursors with multiple copies of the same oligopeptide may have wide occurrence. The precursor to mammalian enkephalin contains six Met-enkephalin sequences and one Leu-enkephalin sequence interspersed within the 267 amino acid precursor (Kilpatrick et al., 1981; Comb et al., 1982). The enkephalins are not as regularly arranged as the four α-factors in the α-factor precursor and are separated by dissimilar oligopeptides. Betz et al. (1981) have found that yeast a cells produce two species of a-factor, both 11 amino acids long, that differ at one position. It would not be surprising to find

Table 4. Strain List

Strain	Relevant Genotype	Source or Reference
DC5	*MATa leu2-3 leu2-112 his3 can1*	J. Hicks; Broach et al. (1979)
DC6	*MATα leu2-3 leu2-112 his4 can1*	J. Hicks
G121C3-24a	*matα2-4 cry1 leu2 his4*	G. Sprague
XK41-10b	*matα2-4 cry1 HMLa leu2-3 leu2-112*	segregant from DC5 × G121C3-24a
RC629	*MATa sst1-2 rme ade2-1 his6*	R. Chan, Chan and Otte (1982)
RC631	*MATa sst2-1 rme ade2-1 his6*	R. Chan, Chan and Otte (1982)
AB320	*HO ade2-1 lys2-1 trp5-2 leu1-12*	K. Nasmyth, Nasmyth and Reed (1980)
G116-4A	*matα1-5 cry1 ura3 trp1 his4 leu2*	G. Sprague
XK96A2-6b	*matα1-5 cry1 leu2-3 leu2-112 trp1*	segregant from G116-4A × DC5
G245-24C	*MATa bar1-1 leu2-3 leu2-112 trp1*	G. Sprague
227	*MATa lys1 cry1*	

that a-factor is produced as a precursor that contains both forms of a-factor. It should be possible to clone the a-factor gene (*MFA*) by a method similar to the one we have used for α-factor and determine its nucleotide sequence to test this proposal.

Why do cells produce multiple copies of a hormone or pheromone from a single precursor? Perhaps the secretion pathway has minimum size requirements for transport into secretion vesicles or for other steps (see Steiner et al., 1980). Production of multiple products from a single precursor would also reduce "shipping costs" (as suggested by Irwin Herskowitz, personal communication), that is, result in more efficient secretion of the mature product for each signal sequence and glycosylation event. Another remarkable finding on mammalian polypeptide hormones is that certain precursors are processed to yield different hormones, as in processing of pro-opiomelanocortin to ACTH, α-MSH, β-endorphin and so forth (Mains et al., 1977; Roberts and Herbert, 1977). Although our sequence data show that the α-factor precursor cannot be processed to yield a-factor, it will be of interest to determine whether the precursor gives rise to other peptides or intermediates in processing that have biological activity.

Experimental Procedures

Strains and Plasmids

Strains are given in Table 4. The yeast clone bank was constructed by Kim Nasmyth in vector YEp13 by insertion of yeast genomic DNA fragments partially digested by endonuclease Sau 3a as described by Nasmyth and Reed (1980). YEp13 is a derivative of pBR322 containing the yeast *LEU2* gene and a yeast origin of replication from the 2μ plasmid (Broach et al., 1979). Plasmids YEp13 and pBR322-*MATα* (containing the Eco RI fragment of *MATα*; Nasmyth and Tatchell, 1980) were also provided by K. Nasmyth. YEp13-*MATα*, which contains a Hind III subfragment from pBR322-*MATα*, was constructed in order to determine optimal conditions for halo formation.

Media

E. coli were grown on LB agar supplemented as necessary with ampicillin (100 μg/ml) or tetracycline (20 μg/ml). Yeast complete medium (YEPD) and synthetic minimal medium (SD) are described in Hicks and Herskowitz (1976b). α-Factor halo assays were performed on YEPD and SD media (BBMB and BBSD, respectively), buffered as described by Fink and Styles (1972).

Enzymes and Radiochemicals

Restriction endonucleases were obtained as follows. Fnu 4H was a gift from D. Russell and M. Smith, Eco RI was obtained from Miles Laboratories and other enzymes were obtained from New England BioLabs. DNA ligase was obtained from P-L Biochemicals, calf alkaline phosphatase (grade I) and E. coli DNA polymerase I (Enzym A:n, Klenow) from Boehringer Mannheim. α-^{32}P-dCTP or α-^{32}P-dATP, used for end-labeling of DNA fragments were obtained from Amersham (2000–3000 Ci/mmole) or from New England Nuclear (600 Ci/mmole or 2000–3000 Ci/mmole).

α-Factor Assays

Production of α-factor by colonies or patches of cells was assayed on agar medium by the halo method (modified from Fink and Styles, 1972) or on thin agar slabs by the confrontation assay (Duntze et al, 1970). For the halo assay, approximately 10^7 a cells were spread onto BBMB or BBSD plates and were imprinted by replica plating with colonies or patches of cells to be tested for α-factor production. In most assays, the a tester was strain RC631, which carries the *sst2-1* mutation (Chan and Otte, 1982). Under these conditions, *matα1-5* and *matα2-4* mutants produce a halo at room temperature but not at 30°C or 34°C. For the confrontation assay, a line of cells (approximately 10^5–10^6) to be tested for α-factor production was streaked on a thin agar slab (either YEPD or minimal medium). Individual a cells (usually *MATa sst1-2* strain RC629) were then placed near the line of cells by micromanipulation and observed for response to α-factor (inhibition of budding, formation of elongated cells). Transformants carrying YEp13-derived plasmids were tested on minimal medium lacking leucine (to select for the plasmid) and were incubated overnight at 30°C before introducing the a tester cells.

Screening Transformants for α-Factor Production

Plasmid DNA was extracted from the E. coli–yeast pool and used to transform *leu2 matα2* strain XK41-10b to Leu$^+$ by selection on minimal medium. These colonies were collected and replated on selective medium and then assayed at 30°C for production of α-factor by the halo assay. Optimal conditions for this assay were developed by monitoring halo formation by colonies of strain XK41-10b carrying plasmid YEp13-*MATα*. Maximal halo size was obtained under the following conditions. α-Factor assays were performed on minimal medium to select for maintenance of YEp13. The tester strains used for the α-factor halo assay carried a mutation in the *SST2* gene, which leads to supersensitivity to α-factor (Chan and Otte, 1982). In addition to transformant K69, a few other transformants were identified that produced smaller halos. These plasmids were not studied further.

Construction of Plasmid Derivatives

The Hind III fragment containing the *MATα* locus was transferred from pBR322-*MATα* to YEp13 by cleavage of pBR322-*MATα* with Hind III and ligating to Hind III–cleaved YEp13 that had been treated with calf alkaline phosphatase. YEp13 carrying the 1.3 kb Hind III fragment of p69A (plasmid YEp13-H1) was constructed in a similar manner. A derivative of p69A lacking the H1 fragment (plasmid YEp13-H2) was obtained by cleavage of p69A with Hind III followed by circularization. Derivatives of p69A lacking different Eco RI fragments were constructed by partial digestion of p69A with Eco RI followed by ligation. Amp^r E. coli transformants that complement the bacterial *leuB* mutation must contain the 4.6, 4.0 and 2.3 kb Eco RI fragments (which includes a small fragment of the yeast insert in p69A, fragment R1-4). Presence of additional Eco RI fragments (R1-1, 1.1 kb; R1-2, 1.7 kb; R1-3, 1.0 kb) was determined by restriction endonuclease digests. Plasmids were then transformed into the *matα2-4* strain and tested for ability to produce α-factor.

DNA Sequencing and Hybridization

The sequencing procedure of Maxam and Gilbert (1977; 1980) was used. DNA fragments were labeled by the 3'-end-labeling technique (Smith et al., 1979). All restriction endonuclease sites used for labeling DNA ends have been read across in an independent sequence analysis. Hybridizations to restriction endonuclease fragments were performed by the method of Southern (1975).

Acknowledgments

We thank Kim Nasmyth for supplying us with his clone bank and other plasmids, Russell Chan and George Sprague for strains, John Yokum for the R-loop analysis, George Sprague for suggestions, Caroline Astell and Michael Smith for discussions on sequencing, Gunther Kreil for calling our attention to processing of melittin and for his suggestions, David Julius and Jeremy Thorner for communicating unpublished information on our *MFα* clone and *ste13* mutant, Michael Smith, Benjamin Hall, Edward Herbert, Irwin Herskowitz, Randy Schekman and Elizabeth Jones for discussion and comments on the manuscript and Flora Banuett for acronymic insight. In addition, we thank Kerrie Rine for preparation of the figures and Virginia Staubach for preparation of the manuscript. This work was supported by a Research Career Development Award and by a research grant from the U. S. Public Health Service to I. H., and by a Damon Runyon–Walter Winchell Cancer Fund fellowship to J. K.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 2, 1982; revised August 2, 1982

References

Astell, C. R., Ahlstrom-Jonasson, L., Smith, M., Tatchell, K., Nasmyth, K. A. and B. D. Hall (1981). The sequence of the DNAs coding for the mating-type loci of Saccharomyces cerevisiae. Cell *27*, 15–23.

Bennetzen, J. L. and Hall, B. D. (1982a). The primary structure of the *Saccharomyces cerevisiae* gene for alcohol dehydrogenase I. J. Biol. Chem. *257*, 3018–3025.

Bennetzen, J. L. and Hall, B. D. (1982b). Codon selection in yeast. J. Biol. Chem. *257*: 3026–3031.

Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980). The ovalbumin gene: sequence of putative control regions. Nucl. Acids. Res. *8*, 127–142.

Betz, R., Duntze, W. and Manney, T. R. (1981). Hormonal control of gametogenesis in the yeast *Saccharomyces cerevisiae*. Gamete Research *4*, 571–584.

Blobel, G. and Dobberstein, B. (1975). Transfer of proteins across membranes: I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J. Cell. Biol. *67*, 835–851.

Broach, J. R. and Hicks, J. B. (1980). Replication and recombination functions associated with the yeast plasmid, 2μ circle. Cell *21*, 501–508.

Broach, J. R., Strathern, J. N. and Hicks, J. B. (1979). Transformation in yeast: development of a hybrid cloning vector and isolation of the *CAN1* gene. Gene *8*, 121–133.

Chan, R. K. and Otte, C. A. (1982). Isolation and genetic analysis of *Saccharomyces cerevisiae* mutants supersensitive to G1 arrest by a-factor and α-factor pheromones. Mol. Cell Biol. *2*, 11–20.

Comb, M., Seeburg, P. H., Adelman, J., Elden, L. and Herbert, E. (1982). Primary structure of the human met- and leu-enkephalin precursor and its mRNA. Nature *295*, 663–666.

Duntze, W., MacKay, V. and Manney, T. R. (1970). *Saccharomyces cerevisiae*; a diffusible sex factor. Science *168*, 1472–1473.

Fink, G. R. and Styles, C. A. (1972). Curing of a killer factor in *Saccharomyces cerevisiae*. Proc. Nat. Acad. Sci. USA *69*, 2846–2849.

Hayashi, R., Moore, S. and Stein, W. (1973). Carboxypeptidase from yeast. Large scale preparation and the application of COOH-terminal analysis of peptides and proteins. J. Biol. Chem. *248*, 2296–2302.

Hicks, J. B. and Herskowitz, I. (1976a). Evidence for a new diffusible element of mating pheromones in yeast. Nature *260*, 246–248.

Hicks, J. B. and Herskowitz, I. (1976b). Interconversion of yeast mating types. I. Direct observation of the action of the homothallism (*HO*) gene. Genetics *83*, 245–258.

Hook, V. Y. H., Elden, L. E. and Brownstein, M. J. (1982). A carboxypeptidase processing enzyme for enkephalin precursors. Nature *295*, 341–342.

Kilpatrick, D. L., Taniguchi, T., Jones, B. N., Stern, A. F., Shively, J. E., Hullihan, J., Kimura, S., Stein, S. and Udenfriend, S. (1981). A highly potent 5200-dalton adrenal opioid peptide that contains both a [Met]- and a [Leu]enkephalin sequence. Proc. Nat. Acad. Sci. USA *78*, 3265–3268.

Kreil, G. (1981) Transfer of proteins across membranes. Ann. Rev. Biochem. *50*, 317–348.

Kreil, G., Haiml, L. and Suchanek, G. (1980). Stepwise cleavage of the pro part of promelittin by dipeptidyipeptidase IV. Eur. J. Biochem. *111*, 49–58.

Leibowitz, M. J. and Wickner, R. B. (1976). A chromosomal gene required for killer plasmid expression, mating, and spore maturation in *Saccharomyces cerevisiae*. Proc. Nat. Acad. Sci. USA *73*, 2061–2065.

MacKay, V. L. and Manney, T. R. (1974). Mutations affecting sexual conjugation and related processes in *Saccharomyces cerevisiae*. I. Isolation and phenotypic characterization of nonmating mutants. Genetics *76*, 255–271.

Mains, R. E., Eipper, B. A. and Ling, H. (1977). Common precursor to corticotropins and endorphins. Proc. Nat. Acad. Sci. USA *74*, 3014–3018.

Manney, T. R., Duntze, W. and Betz, R. (1981). The isolation, characterization and physiological effects of the *Saccharomyces cerevisiae* sex pheromones. In Sexual Interactions in Eukaryotic Microbes, D. O. O'Day and P. A. Horgen, eds. (New York: Academic Press), pp. 21–51.

Maxam, A. M. and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Nat. Acad. Sci. USA *74*, 560–564.

Maxam, A. M. and Gilbert, W. (1980). Sequencing end-labelled DNA with base specific chemical cleavages. Meth. Enzymol. *65*, 499–560.

Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang, A. C. Y., Cohen, S. N. and S. Numa (1979). Nucleotide sequence of cloned cDNA for bovine corticotropin-β-lipotropin precursor. Nature *278*, 423–427.

Nasmyth, K. A. and Reed, S. I. (1980). Isolation of genes by complementation in yeast: molecular cloning of a cell-cycle gene. Proc. Nat. Acad. Sci. USA *77*, 2119–2123.

Nasmyth, K. A. and Tatchell, K. (1980). The structure of transposable yeast mating type loci. Cell *19*, 753–764.

Roberts, J. L. and Herbert, E. (1977). Characterization of a common precursor to corticotropin and β-lipotropin: cell-free synthesis of the precursor and identification of corticotropin peptides in the molecule. Proc. Nat. Acad. Sci. USA *74*, 4826–4830.

Scherer, G., Haag, G. and Duntze, W. (1974). Mechanism of α-factor biosynthesis in *Saccharomyces cerevisiae*. J. Bacteriol. *119*, 386–393.

Smith, M., Leung, D. W., Gillam, S., Astell, C. R., Montgomery, D. L. and Hall, B. D. (1979). Sequence of the gene for iso-1-cytochrome *c* in *Saccharomyces cerevisiae*. Cell *16*, 753–761.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98*, 503–517.

Sprague, G. F. Jr., and Herskowitz, I. (1981). Control of yeast cell type by the mating type locus. I. Identification and control of expression of the a-specific gene, *BAR1*. J. Mol. Biol. *153*, 305–321.

Sprague, G. F., Jr., Rine, J. and Herskowitz, I. (1981). Control of yeast cell type by the mating type locus. II. Genetic interactions between *MATα* and unlinked α-specific *STE* genes. J. Mol. Biol. *153*, 323–335.

Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. and Trager, H. S. (1980). Processing mechanisms in the biosynthesis of proteins. Ann. N.Y. Acad. Sci. *343*, 1–16.

Stern, A. S., Jones, B. N., Shively, J. E., Stein, S. and Udenfriend, S. (1981). Two adrenal opioid polypeptides: proposed intermediates in the processing of proenkephalin. Proc. Nat. Acad. Sci. USA *78*, 1962–1966.

Strathern, J., Hicks, J. and Herskowitz, I. (1981). Control of cell type in yeast by the mating type locus: the α1-α2 hypothesis. J. Mol. Biol. *147*, 357–372.

Struck, D. K., Lennarz, W. J. and Brew, K. (1978). Primary structural requirements for enzymatic formation of the N-glycosidic bond in glycoproteins. J. Biol. Chem. *253*, 5786–5794.

Suarez Rendueles, M. P., Schwencke, J., Garcia Alvarez, N. and Gascon, S. (1981). A new X-prolyl-dipeptidyl aminopeptidase from yeast associated with a particulate fraction. FEBS Lett. *131*, 296–300.

Tanaka, T. and Kita, H. (1977). Degradation of mating factor by α-mating type cells of *Saccharomyces cerevisiae*. J. Biochem. *82*, 1689–1693.

Zaret, K. S. and Sherman, F. (1982). DNA sequence required for efficient transcription termination in yeast. Cell *28*, 563–573.

JONES DAY 3.0-001 FWC

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re patent application of: Gautvik, et al

Serial No.

Filed: Herewith

For: PRODUCTION OF HUMAN PARATHYROID HORMONE FROM MICROORGANISMS

Art Unit 161

Examiner

Hon. Comm. Patents & Trademarks
U.S. Patent & Trademark Office
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Sir:

Preliminary to the prosecution of this application, applicants respectively request that the following amendments be entered and remarks be considered.

AMENDMENTS

IN THE CLAIMS:

Please cancel claims 1, 6, 7, 18, 24, 25, 30 and 31.

Please amend claims 2, 3, 4 and 5 by deleting in the first line of each claim "1" and substituting therefore --43--.

Please amend claims 20 and 21 by deleting in the first line thereof "18" and substituting therefor --43--.

Please amend claim 22 in the second line thereof by deleting "18" and substituting therefor --43--.

Please add the following new claims.

Dup 1

36. A DNA sequence encoding yeast mating factor alpha 1 and human parathyroid hormone wherein said DNA sequence can stably transform a yeast cell thereby conferring on said transformed yeast cell the ability to express and secrete an intact human parathyroid hormone.

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"Express Mail" Mailing Label Number B86979875

Date of Deposit 8-14-89

I hereby certify that this paper and/or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.

(Name)

37. The human parathyroid hormone of claim 36 wherein said hormone is not glycosylated.

38. The DNA sequence of claim 36 wherein said DNA sequence comprises the nucleotide sequence set forth in Figure 11.

39. The yeast cell of claim 36 wherein said cell is in the genus Saccharomyces.

40. The yeast cell of claim 36 wherein said cell is of the species Saccharomyces cerevisiae.

41. The yeast cell of claim 36 wherein said cell is a budding yeast cell.

42. The human parathyroid hormone of claim 36 wherein said hormone has biological activity substantially equivalent to naturally occurring human parathyroid hormone.

43. A plasmid for insertion into yeast and capable of autonomous replication therein comprising the DNA sequence of claim 36.

44. A plasmid for insertion into yeast and capable of autonomous replication therein comprising a DNA sequence encoding the parathyroid hormone of claim 37.

45. A plasmid for insertion into yeast and capable of autonomous replication therein comprising the DNA sequence of claim 38.

46. A plasmid for insertion into yeast and capable of autonomous replication therein capable of transforming the yeast cell of claim 39.

47. A plasmid for insertion into yeast and capable of autonomous replication therein capable of transforming the yeast cell of claim 40.

48. A plasmid for insertion into yeast and capable of autonomous replication therein capable of transforming the yeast cell of claim 41.

49. A transformed yeast cell comprising a DNA sequence encoding yeast mating factor alpha and a human parathyroid hormone, said cell capable of expressing said DNA and secreting said expressed DNA into an extracellular environment, whereby said secreted, expressed DNA is intact human parathyroid hormone.

50. The transformed cell of claim 49 wherein said DNA sequence encoding yeast mating factor alpha encodes yeast mating factor alpha 1.

51. The transformed cell of claim 49 wherein said human parathyroid hormone is not glycosylated.

24 52. The transformed yeast cell of claim 23 49 wherein said DNA sequence comprises the nucleotide sequence of Figure 11.

25 53. The transformed yeast cell of claim 23 49 wherein said yeast cell is of the genus Saccharomyces.

26 54. The transformed yeast cell of claim 23 49 wherein said yeast cell is of the species Saccharomyces cerevisiae.

27 55. The transformed yeast cell of claim 23 49 wherein said yeast cell is a budding yeast cell.

28 56. The transformed cell of claim 23 49 wherein said expressed human parathyroid hormone has a biological activity substantially equivalent to naturally occurring human parathyroid hormone.

57. A transformed yeast cell capable of expressing and secreting intact human parathyroid hormone.

58. The transformed yeast cell of claim 57 wherein said cell expresses and secretes a nonglycosylated, intact human parathyroid hormone.

59. The transformed yeast cell of claim 57 wherein said expressed and secreted intact human parathyroid hormone

has biological activity substantially equivalent to naturally occurring human parathyroid hormone.

60. The transformed yeast cell of claim 59 wherein said intact human parathyroid hormone is not glycosylated.

REMARKS

At the outset, applicants wish to thank the Examiner for the courtesies she extended to applicants' representatives during an interview on June 22, 1989. During that interview, the Examiner and applicants' representatives exchanged their views with respect to the instant file wrapper continuation application. The Examiner advised that declarations setting forth the unexpectedness of applicants' expression and secretion of human parathyroid hormone in yeast would be valued by the Examiner during her examination of any claims in the instant file wrapper continuation case. Accordingly, applicants enclose herewith declarations from two of the inventors, Dr. Karre M. Gautvik ("Gautvik Decl.") and Dr. Odd S. Gabrielson ("Gabrielson Decl."), and a declaration from Dr. Janet Kurjan ("Kurjan Decl."), who is not an inventor in this case.

Applicants have carefully studied the position of the Examiner as set forth in the Official Action from the parent case dated February 14, 1989. In addition, applicants have carefully evaluated the statements made by the Examiner during the aforementioned June 22, 1989 interview. In light of these activities, applicants have filed a file wrapper continuation application concurrently with this Preliminary Amendment. It is respectfully asserted that the amendments made herein, when considered in light of the remarks made herein, render the claims of this continuation case allowable.

Applicants have cancelled claims 1, 6, 7, 18, 24, 25, 30 and 31 from the parent case. The dependency of claims 2, 3, 4, 5, 20, 21 and 22 have been amended so that these claims now depend from new claim 43. New claims designated 36-60 have been added. Of these, claims 36, 49 and 57 are independent claims. Claims 37-48 depend from claim 36, claims 50-56 depend from claim 48 and claims 58-60 depend from claim 57.

All the new claims are limited in some way to the expression and secretion of intact human parathyroid hormone. In this regard, applicants most fervently point out that they are the first to express intact human parathyroid hormone in yeast such that the intact hormone is secreted into an extracellular environment.

To effectuate the processing and secretion of intact human parathyroid hormone, applicants were faced with numerous problems, as set forth in the declarations filed herewith, including the unpredictability of biological expression systems. Gabrielson Decl. ¶ 4. These difficulties were exacerbated by the fact that human parathyroid hormone is an extremely unstable polypeptide. Gautvik Decl. ¶ 4. The instability of the polypeptide is manifest by its ready susceptibility to naturally occurring protease activity including naturally occurring protease activity in yeast cells. Evidence of this protease susceptibility is manifest in <u>in situ</u> biological systems by the large amount of parathyroid hormone fragments found in thyroid glands, e.g., large numbers of carboxy terminal fragments. One aspect of this degradation problem was transport of the expressed human parathyroid hormone to the yeast cell membrane in such a manner so that the hormone is secreted as an intact and properly processed hormone and so that the hormone is not degraded by the aforementioned susceptibility of the hormone to protease attack. Applicants

were both surprised and delighted by their success in overcoming these problems.

The expression system which applicants devised permitted the expression of mature, intact human parathyroid hormone in yeast cell systems by construction of transformed cells, DNA sequences and plasmids that combine the structural DNA sequence for human parathyroid hormone with nonstructural sequences instructing yeast cells to express and secrete the structural hormonal polypeptide. The selection of an appropriate nonstructural leader sequence encoding expression and secretion of the structural sequence and the processing off of this leader sequence without any degradation or destructive effect on the human parathyroid hormone itself is an important aspect of this invention.

In the preferred embodiment of the instant invention, applicants selected the prepro portion of mating factor alpha. It is important to the proper functioning of this embodiment that this nonstructural leader sequence gene contains both the pre portion and the pro portion. Gautvik Decl. ¶ 5. This is contrary to the teachings that suggest that only the pre portion is needed for expression and secretion of polypeptides and yeast. Gabrielson Decl. ¶ 3. Applicants' studies suggest that both the pre and pro portions of mating factor alpha are needed to ensure expression and secretion of intact human parathyroid hormone. Accordingly and as set forth in the instant application, e.g., at page 8, line 25, applicants have used in their preferred embodiment the prepro part of the mating factor alpha 1 gene (see also Figure 11). In light of this aspect of the preferred embodiment, applicants have amended all of the claims from the parent case and submitted new claims in the instant amendment. These new claims all contain a limitation to the aforementioned yeast mating factor

alpha gene which, as shown by the specification, includes the prepro portion thereof.

Respecting the new claims, the Examiner's attention is respectfully directed to independent claims 36, 49 and 57. As is plainly evident from the claims, all of the claims recite as a limitation the requirement of yeast mating factor alpha including the prepro portion and/or secretion thereof of intact hormone. In addition, the claims recite limitations such that the expression of the DNA encoding human parathyroid hormone manifests itself in the expression and secretion of intact human parathyroid hormone.

It is most strongly asserted that these new claims, with the aforementioned limitations, are patentable and manifestly distinguish over the prior art. As applicants have already most vigorously asserted, they are the first to express intact human parathyroid hormone in a yeast system and the first to show significant biologic activity of the expression product. Gautvik Decl. ¶ 6. Applicants' new claims now manifestly recite the expression of such a hormone and the fact that the expressed hormone is intact human parathyroid hormone that is secreted. Applicants have already set forth above that all of the claims are limited in addition to the aforementioned factors the requirement of a yeast mating factor alpha 1 including the prepro region thereof. None of the prior art references of record in the parent case teach this invention. Gautvik Decl. ¶¶ 4, 6, 8 and 11; Kurjan Decl. ¶ 4.

The Kurjan, et al reference, U.S. Patent No. 4,546,082, has been cited by the Examiner for teaching "a general method applicable to obtaining secretion of heterologous polypeptides in Saccharomyces." It is respectfully contended that this statement vastly overstates the teaching of the Kurjan patent. Kurjan Decl. ¶ 4. The

Kurjan patent discloses an important tool for use by molecular biologists, to wit, use of the mating factor alpha pheromone gene for expression of mammalian peptides in yeast cell systems. Kurjan Decl. ¶ 3. In particular, the Kurjan patent uses by way of example three peptides somatostatin, enkaphelin and ACTH. Although the Kurjan work is certainly important, it is not a "cook book" teaching for expression of mammalian peptides in yeast. Gautvik Decl. ¶ 9. For example, reports are extant indicating a lack of success using the mating factor alpha gene in the expression, processing and secretion of heterologous peptides. Kurjan Decl. ¶ 4. This failure of others is in part due to the unpredictability of biological systems, Kurjan Decl. ¶ 5, and indicates that it would not be at all expected that the Kurjan work could be applied to expression of intact, properly processed human parathyroid hormone. Gabrielson Decl. ¶¶ 5, 6. In view of the fragility of human parathyroid hormone and the established unpredictability of using mating factor alpha in genetic constructions for expression of heterologous peptides, the accomplishment of this invention is truly noteworthy. Kurjan Decl. ¶ 8; Gabrielson Decl. ¶ 8.

One of the reasons the Examiner has overstated the teaching of Kurjan is the Examiner's apparent failure to appreciate the nature of proteolitic degradation and proteolitic processing in yeast cells. Yeast cells are quite rich is protease activity. Gabrielson Decl. ¶ 4. In addition, the amino acid sequence of human parathyroid hormone suggests that it would be susceptible to cleavage by the naturally occurring yeast protease, KEX-2 endopeptidase. This is a significant problem for the inventors to overcome because the KEX-2 endopeptidase is the protease responsible for cleavage of the pro region of mating factor alpha. Kurjan Decl. ¶ 7.

With respect to the Mahoney reference cited in the parent case, applicants respectfully but most pointedly note that only a miniscule portion of the hormone cloned in Mahoney is secreted outside the cell. Particularly, with reference to the table on Page 20 of the Mahoney published European application, it is evident that only approximately one-ninth of the expressed hormone is outside the cell and, of that tiny amount, virtually no biologic activity is seen outside the cell. Gautvik Decl. ¶ 10. In addition, the radiommunoassay used to detect the presence of the hormone extracellularly did not assay for complete hormone but for one of three fragments thereof. Accordingly, it is respectfully urged that Mahoney explicitly demonstrates the great difficulty in obtaining secretion of the human parathyroid hormone outside a cell in any appreciable amount and the extreme difficulty in obtaining such secretion in a manner so as to preserve biological activity. Moreover, Mahoney uses a different genetic construction than the instant invention. Kurjan Decl. ¶ 6. Thus, the achievements of applicants are highly noteworthy and certainly nonobvious in view of the failure of Mahoney, et al.

Kronenberg is clearly distinguished from the present invention. Firstly, the genetic construction of Kronenberg is not the same as that used in the instant invention. Kurjan Decl. ¶ 6. Kronenberg did not use mating factor alpha. Moreover, the expressed PTH was not processed nor was it secreted in the yeast expression system of Kronenberg. *Id.* Accordingly, Kronenberg demonstrates the difficulty of expressing human parathyroid hormone in yeast and, therefore, manifestly evidences the nonobviousness of applicant's invention, *inter alia*, because it does not provide how to make a genetic construction that would permit the proper expression, processing and secretion of intact human parathyroid hormone.

Kurjan Decl. ¶ 8. Indeed, the difficulty of achieving the proper genetic construction is evidenced by the failure of applicants to obtain proper expression, processing and secretion of human parathyroid hormone by their own failures. Gabrielson Decl. ¶ 3. Regarding the Examiner's contention that the failure of Kronenberg was obvious because Kronenberg failed to use a yeast signal sequence, applicants refer the Examiner to the statements made by Dr. Kurjan, the discoverer of the benefits of the yeast mating factor alpha system in genetic constructions, that the use of this genetic tool is not a guarantor that all mammalian peptides can effectively be expressed, processed and secreted in yeast cells using the mating factor alpha system. Kurjan Decl. ¶ 4.

It is important to appreciate the unique juxtaposition of problems faced by the applicants in their construction of this invention. Applicants were faced with the dilemma of using the proteolytic enzymes of the yeast cell for cleavage of the mating factor alpha leader sequence from the downstream heterologous protein, human parathyroid hormone, while at the same time ensuring that this very proteolytic activity required for cleavage of mating factor alpha did not cleave, degrade or adversely affect human parathyroid hormone following its expression and secretion from a yeast host cell. Kurjan Decl. ¶ 7. The proteolytic enzyme is believed to cleave mating factor alpha at its C-terminal end is KEX-2 endopeptidase. This endopeptidase acts on a substrate comprising at least two basic amino acids. The structure of human parathyroid hormone is such that it contains at least two distinct such basic amino acid regions making both these regions potential cleavage cites for the endopeptidase enzyme. Gautvik Decl. ¶ 7. Accordingly, one of the hurdles that the applicants overcame in their invention was to exploit the

proteolytic activity of endopeptidase for cleavage of mating factor alpha without having this very same endopeptidase attack basic amino acid cleavage cites extant in human parathyroid hormone. Gabrielson Decl. ¶ 6. Applicants most forcefully contend that no other researchers or inventors faced with this dual problem were successfully able to achieve expression and secretion of hormones such as human parathyroid hormone using yeast mating factor alpha leader sequences. In this regard, attention is directed to the declaration of Dr. Gabrielson at paragraph 7 thereof in which the fact that certain heterologous peptides have been expressed in yeast is discussed but, as Dr. Gabrielson points out, none of these peptides has the proteolytic susceptibility of human parathyroid hormone. Indeed, as attested to by Dr. Gabrielson, those peptides having the proteolytic susceptibility of human parathyroid hormone have either not been expressed or not been secreted using the alpha mating factor system. Gabrielson Decl. ¶ 7. In view of this applicants respectfully contend that none of the references from the parent case so much as suggest that applicants could have been successful in their achievement of expressing and secreting human parathyroid hormone.

Applicants believe that the other references cited by the Examiner in the final action of the parent case are considerably less pertinent to the currently claimed invention when compared with the aforediscussed three references of Kurjan, Mahoney and Kronenberg. Gautvik Decl. ¶ 11. Nonetheless, applicants would take this opportunity to briefly point out some of the distinguishing features between their invention and these less pertinent references. In Breyel, et al, the expression product is formylmethionyl human parathyroid hormone which is expressed only in the cytoplasm and only in extremely small quantities. Born, et al uses the human prepro

sequences of human parathyroid hormone (not the prepro portion of mating factor alpha) to transform the whole cell. Moreover, no production of intact human parathyroid hormone is obtained. Rather, fragments of the human parathyroid hormone are obtained. Finally with respect to the Born reference, there is absolutely no data or suggestion that intact parathyroid hormone is obtained at all, let alone intact human parathyroid hormone that has been secreted in a biologically active form as is the case in the instant invention. Applicants note, in closing this portion of the remarks, that the remaining references cited by the Examiner, Gonoza, et al and Chang, et al, are cited only for extremely general teachings. Gonoza teaches the effect of nucleotides flanking the E. coli start codon and Chang teaches that translational efficiency of eukaryotic proteins in prokaryotes is a function of the distance between the start codon and the ribosome binding site. It is noted that all the now pending claims are limited to expression of human parathyroid hormone in a eukaryotic cell, namely, yeast.

In view of the foregoing, applicants strongly hold the opinion that all of the claims set forth in this Preliminary Amendment are in condition for allowance. Accordingly, applicants respectfully urge the Examiner to allow the claims.

No fee is deemed necessary for this Preliminary Amendment. However, if the Examiner believes a fee is due, she is authorized to charge Deposit Account No. 12-1095. If the Examiner has any inquiries with respect to this case, she is encouraged to contact applicants' counsel.

Respectfully submitted,

LERNER, DAVID, LITTENBERG,
KRUMHOLZ & MENTLIK

By Bruce H. Sales
Bruce H. Sales
Reg. No. 32,793

600 South Avenue West
Westfield, New Jersey 07090
(201) 654-5000